ATTORNEY'S DOCKET NO. SGL-001

PATENTS

UNITED STATES PATENT APPLICATION

OF

SIMSON L. GARFINKEL

FOR

SYSTEM AND METHOD THAT PROVIDES FOR THE EFFICIENT AND EFFECTIVE SANITIZING OF DISK STORAGE UNITS AND THE LIKE

Certificate of Express Mailing

Express Mail Mailing Label No. EK 904 503 507 US

Date of Deposit August 9, 2001

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office To Addressee" Service under 37 CFR 1.10 on the date indicated above and is addressed to the Commissioner of Patents and Trademarks, Washington, D. C. 20231.

By Richard Jordan

Richard A. Jordan

FIELD OF THE INVENTION

The invention relates generally to the field of digital data processing and storage, and more specifically to systems and methods that provide for sanitizing of disk storage units and the like.

BACKGROUND OF THE INVENTION

Digital computers, mass digital data storage subsystems and the like typically include disk storage units to provide for relatively long-term storage of digital data. It is often necessary to remove a disk storage unit from a computer, mass storage subsystem, etc., in which it formed component. This may occur, for example, if it is necessary to remove the disk storage unit for repair, if it is necessary to replace the disk storage unit, if the computer, mass storage subsystem, etc., is to be discarded, or for other reasons that will be apparent to those skilled in the art.

Data stored in a disk storage unit is often confidential to the organization that maintains the computer, mass storage subsystem, etc., in which the disk storage unit forms a component. Several problems can arise in connection with maintaining the confidentiality of the data that is stored in a disk storage unit. For example, disk storage units store data, organized into files, in magnetic form. Typically, when a file is deleted, the data is not erased from the disk storage unit, but instead information detailing the locations of the data comprising the respective file is deleted from tables that are maintained therefor by the computer, mass storage subsystem, etc. Accordingly, merely

erasing files from a disk storage unit will not serve to erase the data contained in the files. The data can be recovered using any of a number of conventional data recovery techniques.

Even if efforts are made to sanitize a disk storage unit, that is, to erase the data stored in a disk storage unit, to over-write the data with other data, or to perform other sanitizing operations that will be apparent to those skilled in the art, it is often still possible to recover the erased or over-written data, since it is not unusual for residual magnetic fields to remain after the erasure or overwriting that can be detected sufficiently for the data represented thereby to be reconstructed. In addition, since the data storage capacity of disk storage units is quite large and growing, the time required to over-write the data stored on a typical disk storage unit even once is prohibitive, and typically data is not considered "wiped" until it has been over-written at least several times, generally with predetermined data patterns.

Moreover, during wiping, a "Trojan horse" program can cause data to be copied from the storage locations in which it is currently stored to spare storage locations on the disk storage unit that may be provided to accommodate the possibility that some of the "regular" storage locations may go bad. If a regular storage location does go bad, the disk storage unit automatically stores the data that is to be stored on the bad regular storage location on a spare storage location that has been allocated therefor. Thereafter, when the data is to be retrieved from a "regular" storage location for which a spare storage location has been allocated, the disk storage unit will automatically retrieve the data from the spare storage location and provide the data to the device that requested the data. Generally, the spare storage locations will be known to the disk storage unit, and not to the device, that is, the computer or the like that stores data in, and retrieves data from the disk storage unit, and

so the wiping will be in connection with the regular storage locations and not the spare storage locations. In that case, the data will still be available in the spare storage locations.

Instead of overwriting or wiping a disk storage unit, the contents of a disk storage unit can be erased in a "bulk erasure" operation by bringing the disk storage unit in close proximity to a strong magnetic field to "de-gauss" the disk storage unit. However, de-gaussing a disk storage unit, in addition to erasing the data stored thereon, will also erase formatting information that identifies the storage locations, making the disk storage unit thereafter unusable.

SUMMARY OF THE INVENTION

The invention provides a new and improved system and method for the efficient and effective sanitizing of disk storage units, while additionally providing that the disk storage units can be subsequently used without risking the confidentiality of data previously stored thereon.

In brief summary, generally the invention provides for sanitizing of a digital data storage unit, such as a disk data storage unit, by encrypting the information that is stored thereon. When the previously-stored information is retrieved from the digital data storage system storage unit, the information is decrypted prior to being provided to the device that requested retrieval of the data. If the digital data storage unit is to be sanitized, the key or keys that used to at least decrypt the data stored on the digital data storage unit are discarded or made unavailable for use in decrypting the encrypted digital data, thereby making the unencrypted data unavailable. Any of a number of types of encryption/decryption methodologies can be used, including a symmetric key methodology, an

asymmetric key methodology such as a public key/private key methodology, or any of a number of other encryption/decryption methodologies as will be apparent to those skilled in the art.

The decryption key can be provided to the disk storage unit in any of a number of ways that would facilitate discarding of the decryption key when the disk storage unit is to be sanitized. For example, the decryption key may be stored in non-volatile random-access memory ("NVRAM") and available for use in decrypting data retrieved from the disk storage unit. When the disk storage unit is to be sanitized, the NVRAM can be made unavailable to the person or entity that has possession of the disk storage unit, by, for example, discarding the NVRAM, erasing the NVRAM, destroying the NVRAM, or any other arrangement whereby the contents of the NVRAM are unavailable to the person or entity that has possession of the disk storage unit. Since this makes the decryption key unavailable to the person or entity that has possession of the disk storage unit, the person or entity that has possession of the disk storage unit will be unable to retrieve the data in unencrypted form.

Alternatively, some or all of the decryption key can be stored on the disk storage unit itself, and another portion provided in, for example, an integrated circuit ("IC") chip, and the key obtained by performing a selected processing operation in connection with the portion of the key stored on the disk storage unit and the portion stored on the IC chip. For example, the decryption key can be formed by concatenating the portion stored on the IC chip with the portion stored on the disk storage unit. Alternatively, the decryption key can be formed by performing an exclusive-OR ("XOR") operation in connection with the portion stored on the disk storage unit and the portion stored on the IC chip. Other types of processing operations that can be used in connection with the portion stored on the disk storage unit and the portion stored on the IC chip, which can be used to generate a

decryption key, will be apparent to those skilled in the art. When the disk storage unit is to be sanitized, the IC chip can be removed, destroyed, or the like, which can make it impossible to reconstruct the key(s).

BRIEF DESCRIPTION OF THE DRAWINGS

This invention is pointed out with particularity in the appended claims. The above and further advantages of this invention may be better understood by referring to the following description taken in conjunction with the accompanying drawings, in which:

FIG. 1 is a functional block diagram of a mass storage subsystem including a system that provides for the efficient and effective sanitizing of disk storage units included therein, constructed in accordance with the invention; and

FIG. 2 is a functional block diagram of a second embodiment of a mass storage subsystem including a system that provides for the efficient and effective sanitizing of disk storage units included therein, constructed in accordance with the invention.

DETAILED DESCRIPTION OF AN ILLUSTRATIVE EMBODIMENT

FIG. 1 is a functional block diagram of a mass storage subsystem 10 including a system that provides for the efficient and effective sanitizing of disk storage units included therein, constructed in accordance with the invention. With reference to FIG. 1, the mass storage subsystem 10 includes

one or more disk storage units 11(1) through 11(N) (generally identified by reference numeral 11(n)), a controller 12, an interface 13 and one or more key stores 14(1) through 14(N) (generally identified by reference numeral 14(n)). Each disk storage unit 11(n) stores digital data provided to the mass storage subsystem 10 by an external data utilization device (not shown), which digital data can later be retrieved, by the same or a different data utilization device, for usage. The data utilization device(s) may include any of a number of types of devices, including, but not limited to, computers, including personal computers, computer workstations, mini- and mainframe computers, that retrieve data from the mass storage subsystem 10, process the data, and provide processed data to the mass storage subsystem 10 for subsequent storage. In addition, the data utilization device(s) may include any of a number of types of other devices, including data visualization devices that display the data, in text, image or any other form, to a user, devices for producing hardcopy output, network or other communication devices for transferring data over a network or other communication medium to other data utilization devices, backup devices for producing backup copies of data stored on the mass storage subsystem 10, as well as any of a number of other types of devices that can produce and/or make use of digital data.

The interface 13 connects to one or more of the data utilization devices over a communication link 16. The communication link 16 may be any of a number of types of communication links over which information in digital form may be transferred. The interface 13 receives storage retrieval requests from data utilization devices over the communication link 16. A storage request initiates a storage operation in connection with the mass storage subsystem 10, to facilitate the storage of data on one or more of the disk storage units 11(n). Storage requests may

be accompanied by the data to be stored, or the data may be provided to the mass storage subsystem 10 separately from the request. A retrieval request initiates a retrieval operation to enable data to be retrieved from one or more of the disk storage units 11(n). After the data has been retrieved, it can be transferred by the interface 13 over the communication link 16 to the device that issued the retrieval request, or to one or more other devices as directed by the retrieval request.

After the interface 13 has received a storage or retrieval request, it will provide the request to the controller 12 for processing. The controller performs a number of operations in connection with the mass storage subsystem 12, including scheduling of storage and retrieval operations by the respective disk storage units 11(n), buffering of data to be stored in a storage operation pending storage in a disk storage unit 11(n), buffering of data retrieved from a disk storage unit 11(n) prior to transmission to the destination data utilization device by the interface 13, and the like.

In addition, the controller 12 includes an encryption/decryption module 15 that performs an encryption operation in connection with data to be stored on a disk storage unit 11(n) to encrypt the data before it is transferred thereto for storage, and a decryption output in connection with data retrieved from a disk storage unit 11(n) to decrypt the data after it has been retrieved and before it is transferred by the interface 13 to the destination data utilization device. When performing an encryption or decryption operate in connection with data to be stored on or that has been retrieved from a disk storage unit 11(n), the encryption/decryption module 15 makes use of a key stored in a correspondingly-indexed key store 14(n). The encryption/decryption module 15 may make use of any of a number of encryption/decryption methodologies, including a symmetric key methodology, an asymmetric key methodology such as a public key/private key methodology, or any of a number

of other encryption/decryption methodologies as will be apparent to those skilled in the art. If the encryption/decryption module 15 makes use of a symmetric key methodology in connection with a disk storage unit 11(n), the key store 14(n) will store one key that it will use for both encryption of data to be stored and decryption of data that has been retrieved. On the other hand, if the encryption/decryption module makes use of an asymmetric key methodology in connection with a disk storage unit 11(n), the key store 14(n) will store two keys, namely, an encryption key that it will use for encryption of data to be stored and a decryption key that it will use for decryption of data that has been retrieved. It will be appreciated that the encryption/decryption module 15 may make use of different methodologies for different ones of the disk storage units 11(n), in which case one or more of the key stores 14(n) may store one key used for both encryption and decryption, and others may store two keys, one used for encryption and the other used for decryption. In addition, it will be appreciated that data stored on one or more of the disk storage units 11(n) may not be encrypted, in which case no key need be provided therefor.

As noted above, one or more of the disk storage units 11(n) generally store data provided thereto by the controller 12 in encrypted form. For the disk storage units 11(n) for which data is stored in encrypted form, as long as the key is available to decrypt the data stored on the disk storage unit 11(n), the data stored on the disk storage unit 11(n) can be decrypted and provided to a data utilization device. Accordingly, if it is desired to make the data stored on a disk storage unit 11(n) unavailable for access, at least in unencrypted form, the key that is used for decryption is discarded. This can be accomplished in a number of ways. For example, the key store 14(n) can merely be erased, if it the key store 14(n) is in the form of a volatile memory. Alternatively, if the key store

14(n) is in the form of a non-volatile memory, the key store 14(n) can be removed from the mass storage subsystem 10. In that case, erasure or destruction of the non-volatile memory comprising the key store 14(n) can ensure that the data stored on the disk storage unit 11(n) will remain unavailable for access in unencrypted form; the non-volatile memory can be erased using any of a number of conventional erasure techniques, including, for example, application of a selected voltage to the circuitry comprising the key store. As another alternative, the key store 14(n) can comprise a selected storage location on the disk storage unit 11(n) itself, in which case the key can be erased by erasing the selected storage location. As yet another alternative, the key or keys to be used in connection with a disk storage unit 11(n) can initially be stored on the disk storage unit, and thereafter copied by the controller 12 to a separate volatile or non-volatile key store 14(n); in that case, when it is desired to make the data on the disk storage unit 11(n) unavailable in unencrypted form, operations can be performed in a manner similar to those described above in connection with both the disk storage unit 11(n) and the key store 14(n).

If the key or keys used by the encryption/decryption module 15 to encrypt data to be stored on a disk storage unit 11(n), and to decrypt data retrieved from the respective disk storage unit 11(n), are stored in a key store 14(n) separate and apart from the respective disk storage unit 11(n), the key or keys may be initially provided in a number of ways. For example, the key(s) can be initially stored on the disk storage unit 11(n) and copied by the controller 12 to the respective key store 14(n) as part of an initialization procedure when the disk storage unit is first installed in the mass storage subsystem 10, when the mass storage subsystem 10 is first powered on or configured, or as part of any other procedure as will be apparent to those skilled in the art. The controller 12 may also enable

the disk storage unit 11(n) to erase the key(s) from the disk storage unit 11(n) if, for example, the key store 14(n) is a non-volatile memory. Alternatively or in addition, the key(s) may be provided in machine readable form on another machine readable medium that may be read by the controller 12 using a suitable reading device (not separately shown).

As a further alternative, the controller may be provided with a key generator module that can generate the key(s) to be used with the disk storage units. One illustrative embodiment thereof will be described in connection with FIG. 2. FIG. 2 depicts a mass storage subsystem 110 that includes one or more disk storage units 111(n), an interface 113 connected to a communication link 116, key store(s) 114(n) and a controller 112 that generally correspond to and operate in a manner similar to respective disk storage unit(s) 11(n), interface 13, communication link 116, key stores 14(n) and controller 112, except as follows.

In addition to an encryption/decryption module 115, which corresponds to and operates in a manner similar to encryption/decryption module 15, the controller 112 is also provided with a key generator 117 that generates one or more keys for a disk storage unit 111(n) by use of a bit pattern stored on, for example, the respective disk storage unit 111(n) itself and another bit pattern. The bit pattern stored on the respective disk storage unit 111(n) may be stored on, for example, a storage location on the disk storage unit identified as pattern store 119(n). The other bit pattern used by the key generator 117 in generating one or more keys for the disk storage unit 111(n) may be provided in another pattern store 118(n), which may comprise a component of the mass storage subsystem 110, a machine-readable medium that is readable by an appropriate reading device, or the like. The mass storage subsystem 110 may include a single pattern store for storing a bit pattern that can be

used for all disk storage unit(s) included therein, a plurality of pattern stores 118(n) each of which stores a bit pattern for a correspondingly-indexed disk storage unit 111(n), or a plurality of pattern stores each of which stores a bit pattern that will be used in connection with generating one or more keys for selected ones of the disk storage units 111(n). The key generator 117 can perform any of a number of processing operations in connection with the bit patterns, which may include concatenating the bit patterns together, exclusive-ORing ("XORing") some or all of the bit patterns together, or any other types of processing operations as will be apparent to those skilled in the art.

When it is desired to sanitize a disk storage unit 111(n) in the mass storage subsystem 110, the decryption key can be made unavailable by erasing or destroying the key store 114(n) and erasing or destroying one or both of the pattern stores 118(n), 119(n), thereby to make one or both of the bit patterns stored therein unavailable. Making one of the bit patterns unavailable will generally suffice to make regeneration of the key(s) difficult if not impossible.

The invention provides a number of advantages. In particular, the invention provides an arrangement whereby a disk storage unit can be efficiently and effectively sanitized in such a manner that it can subsequently provided to a repair facility, be used in another environment, and the like, while maintaining the confidentiality of the data stored thereon. This is accomplished by providing that the data stored on the disk storage unit is encrypted, and decrypted during a retrieval operation, and further providing that the key used in decrypting the data be unavailable when it is desired to sanitize the disk storage unit. The key used in decrypting the data can be made unavailable by, for example, erasing or destroying a store in which the key is stored; if the key is stored in multiple stores, preferably all of the stores will be, for example, erased or destroyed so that the key will

thereafter be unavailable for use in decryption. Without the key, the data stored on the disk storage unit generally cannot be retrieved in unencrypted form, in which case the data will remain confidential if the disk storage unit is provided to an entity other than that for which the data comprises confidential information.

It will be appreciated that a number of changes and modifications may be made to the arrangement described herein. As noted above, any of a number of conventional encryption/decryption methodologies may be used, including both symmetric key and asymmetric key methodologies. It will be appreciated that generally the ability of an encryption/decryption methodology to maintain the confidentiality of the encrypted data will reflect the security of the methodology and the number of bits comprising the key(s) used in encryption and decryption. In addition, it will be appreciated that, if it is desired to have the data stored on the disk storage unit available in unencrypted form after the disk storage unit has been sanitized, it will generally be desirable to backup the data before the disk storage unit is sanitized, using any of a number of conventional backup arrangements. The data may be backed up directly onto another disk storage unit, or onto a backup medium such as tape, after which it can be loaded onto another disk storage unit for subsequent retrieval.

Although the arrangement has been described as being used in connection with disk storage units, it will be appreciated that the arrangement can also be used in connection with other types of digital data storage arrangements, including storage subsystems that emulate disk storage units but make use of storage media other than magnetic disks.

It will be appreciated that a system in accordance with the invention can be constructed in whole or in part from special purpose hardware or a general purpose computer system, or any combination thereof, any portion of which may be controlled by a suitable program. Any program may in whole or in part comprise part of or be stored on the system in a conventional manner, or it may in whole or in part be provided in to the system over a network or other mechanism for transferring information in a conventional manner. In addition, it will be appreciated that the system may be operated and/or otherwise controlled by means of information provided by an operator using operator input elements (not shown) which may be connected directly to the system or which may transfer the information to the system over a network or other mechanism for transferring information in a conventional manner.

The foregoing description has been limited to a specific embodiment of this invention. It will be apparent, however, that various variations and modifications may be made to the invention, with the attainment of some or all of the advantages of the invention. It is the object of the appended claims to cover these and such other variations and modifications as come within the true spirit and scope of the invention.

What is claimed as new and desired to be secured by Letters Patent of the United States is: